Matrix Element and Monte Carlo Generation and the Les Houches Accords

J. Huston

IPPP Workshop

Theoretical Predictions for New (Old) Physics

There are a variety of programs available for comparison of data to theory and/or predictions.

- Tree level
- Leading log Monte Carlo
- N^nLO
- Resummed

In general, agree quite well…but before you appeal to new physics, check the ME.
Can have ME corrections to MC or MC corrections to ME(**a la Les Houches acco**

Perhaps biggest effort…include NLO ME corrections in Monte Carlo programs…correct normalizations. Correct shapes.
N^nLO needed for precision physics.

Resummed description describes soft gluon effects (better than MC's)…has correct normalization (but need HO to get it); resummed predictions include non-perturbative effects correctly…may have to be put in by hand in MC's

threshold
dijet, direct γ

k_T
W,Z, Higgs

b space
(ResBos)

q_t space

Important to know strengths/weaknesses of each.

Where possible, normalize to existing data.

…in addition, worry about pdf, fragmentation uncertainties

Les Houches Update

- Two workshops on "Physics at TeV Colliders" have been held so far, in 1999 and 2001 (May 21-June 1)
- Working groups on QCD/SM, Higgs, Beyond Standard Model
- See web page:

http://wwwlapp.in2p3.fr/conferences/LesHouches/Houches2001/

especially for links to writeups from 1999 and 2001

- QCD 1999 writeup (hep-ph/0005114) is an excellent pedagogical review for new students
- QCD 2001 writeup (hep-ph/0204316) is a good treatment of the state of the art for pdfs, NLO calculations, Monte Carlos
- Les Houches 2003 will have more of a concentration on EW/top physics

Monte Carlo Interfaces

- To obtain full predictability for a theoretical calculation, would like to interface to a Monte Carlo program (Herwig, Pythia, Isajet)
  - parton showering (additional jets)
  - hadronization
  - detector simulation
- Some interfaces already exist
  - VECBOS→Herwig (HERPRT)
  - CompHep→Pythia
- A general interface accord was reached at the 2001 Les Houches workshop
- All of the matrix element programs mentioned will output 4-vector and color flow information in such a way as to be universally readable by all Monte Carlo programs
- CompHep, Grace, Madgraph, Alpha, etc, etc

  →Herwig, Pythia, Isajet

Les Houches and Monte Carlos

- Much of the time during meeting was spent developing a generic process interface from matrix element to Monte Carlo programs
- This interface allows:
  - arbitrary hard subprocesses to be plugged into shower/hadronization generators.

    CompHEP
    Grace → Herwig
    MadGraph → Isajet
    VecBos → Pythia
    Wbbgen
  - **->Les Houches accord (#1)**

"Les Houches" User Process Interface for Event Generators

hep-ph/0109068

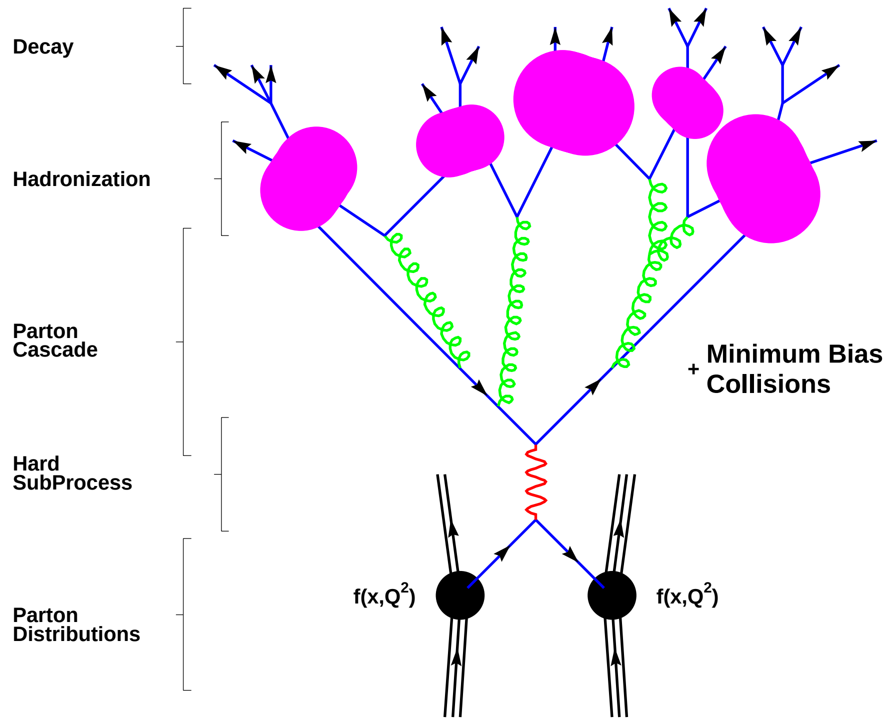
E. Boos, M. Dobbs, W. Giele, I. Hinchliffe, J. Huston,
V. Ilyin, J. Kanzaki, K. Kato, Y. Kurihara,
L. Lönnblad, M. Mangano, S. Mrenna, F. Paige, E. Richter-Was,
M. Seymour, T. Sjöstrand, B. Webber, D. Zeppenfeld

- Possible because one or more authors from each of these programs was present at Les Houches
  - Matt Dobbs has been the front man for coordinating the disputes/discussions
  - literally hundreds of email exchanges

Universal Interface

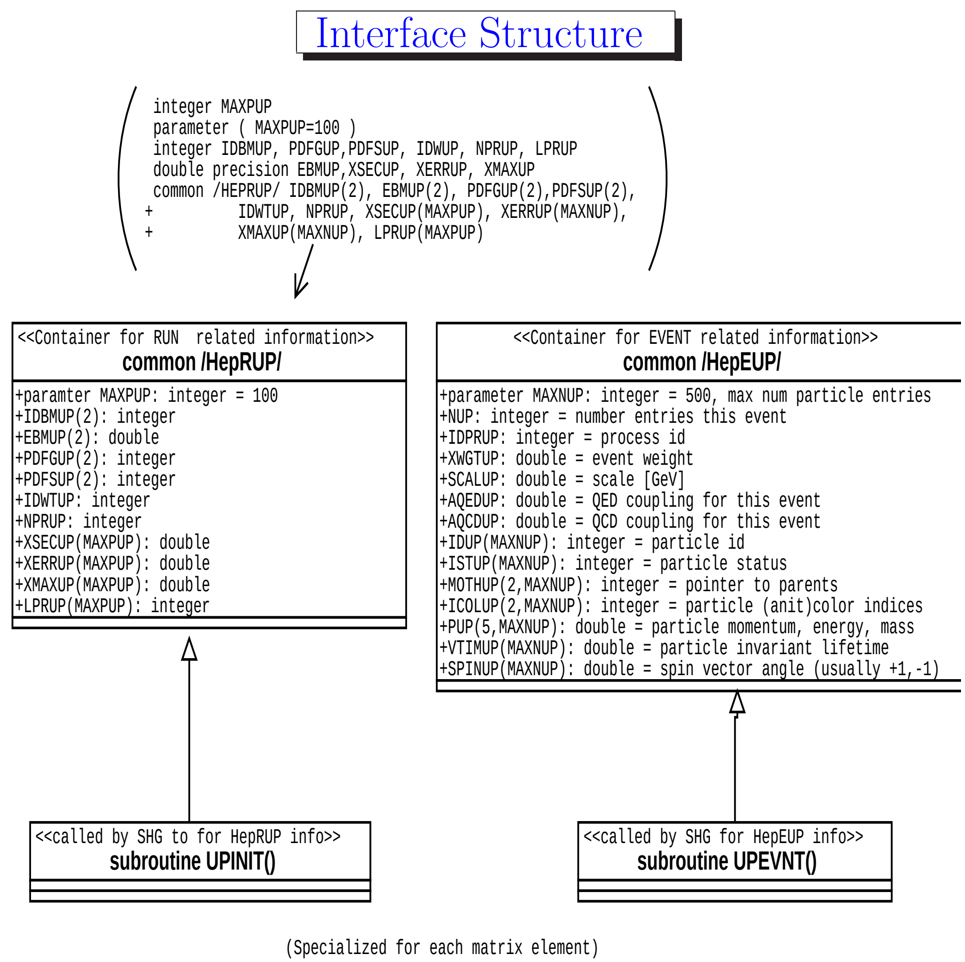
- This interface will allow for a more complete predictability for ME programs
  - parton showering (additional jets)
  - hadronization
  - detector simulation
- Some specialized interfaces already exist
  - VECBOS→Herwig (HERPRT)
  - Wbbgen→Herwig
  - CompHep→Pythia
- This interface should supercede them.

Specialize in the 'generic' parts of the event.



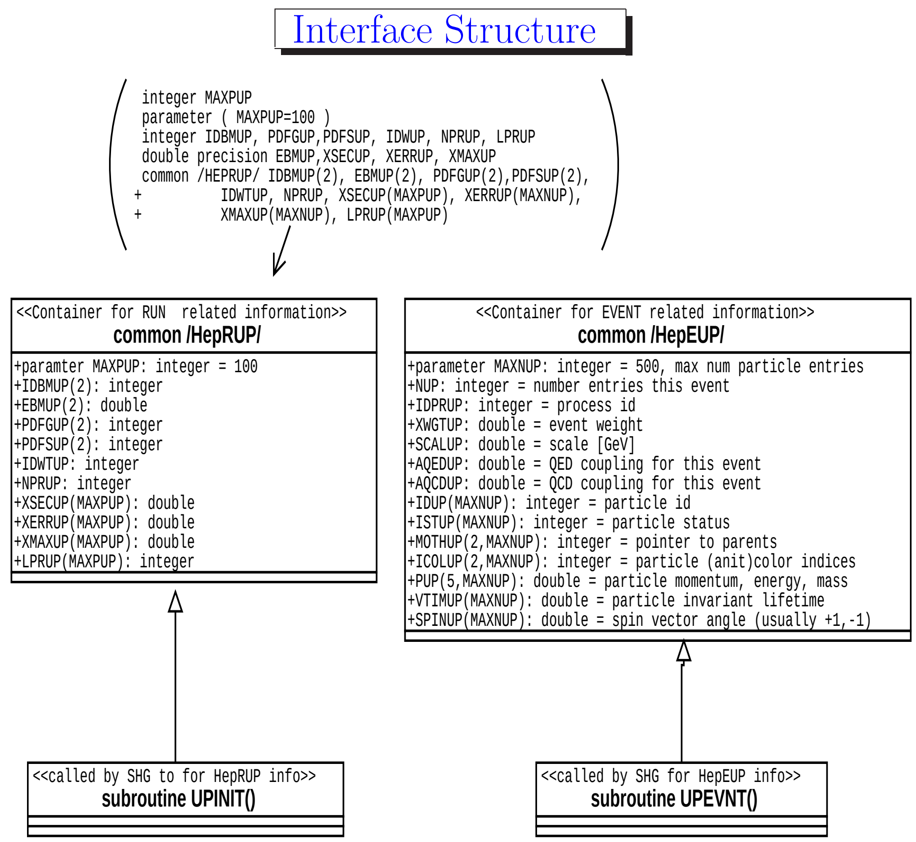
Interface

- Provides information on parton 4-vectors,mother-daughter relationships, spins/helicities and color flow
  - also points to intermediate particles whose mass should be preserved in the parton showering
- Not intended as a replacement for HEPEVT
  - addresses communication between event generators only, not between event generators and the outside world
- Partonic information is in 2 Fortran common blocks
  - run info
  - specific event info



Subroutines

- Each stage (run and event) associated with own subroutine, called from the shower generator, where information is placed in the respective common block, based on output from the matrix element generator
- Subroutine names (in Pythia 6.2) are:
  - UPINIT
  - UPEVNT
  - **note no PY prefixes**
- Other authors should use the same convention



(Specialized for each matrix element)

Unweighting

- Shower generator can unweight events from matrix element generator, mix different subprocesses from matrix element generator, or just read events straight from a file
  - if unweighting/mixing is needed then shower generator needs info about subprocess cross sections and/or maximum weights
- If extra information is needed for specific user implementation, then implementation-specific common block has to be created
- Note that a lot of the technicalities are intended for ME/MC authors, not for users; in most cases, these details will be invisible to the casual user

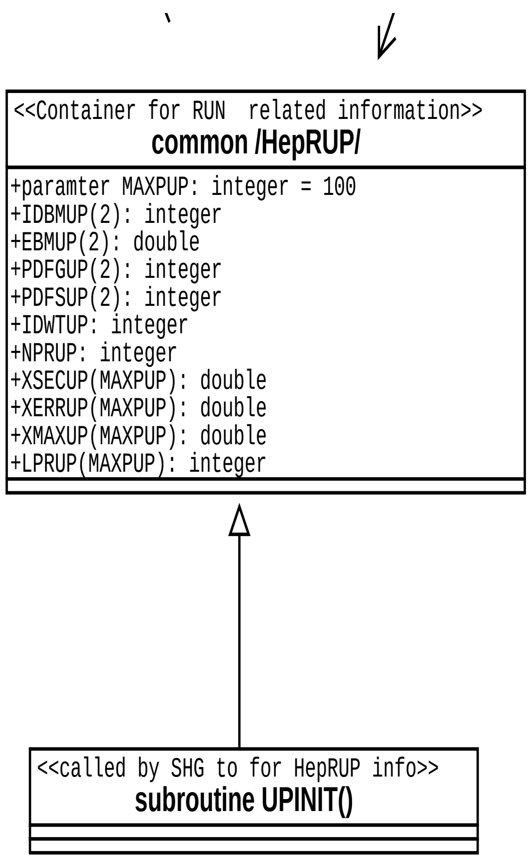
Interface Structure

```
      integer MAXPUP
      parameter ( MAXPUP=100 )
      integer IDBMUP, PDFGUP,PDFSUP, IDWUP, NPRUP, LPRUP
      double precision EBMUP,XSECUP, XERRUP, XMAXUP
      common /HEPRUP/ IDBMUP(2), EBMUP(2), PDFGUP(2),PDFSUP(2),
     +          IDWTUP, NPRUP, XSECUP(MAXPUP), XERRUP(MAXNUP),
     +          XMAXUP(MAXNUP), LPRUP(MAXPUP)
```

MAXUP: maximum number of different processes to be interfaced at one time

Run related information

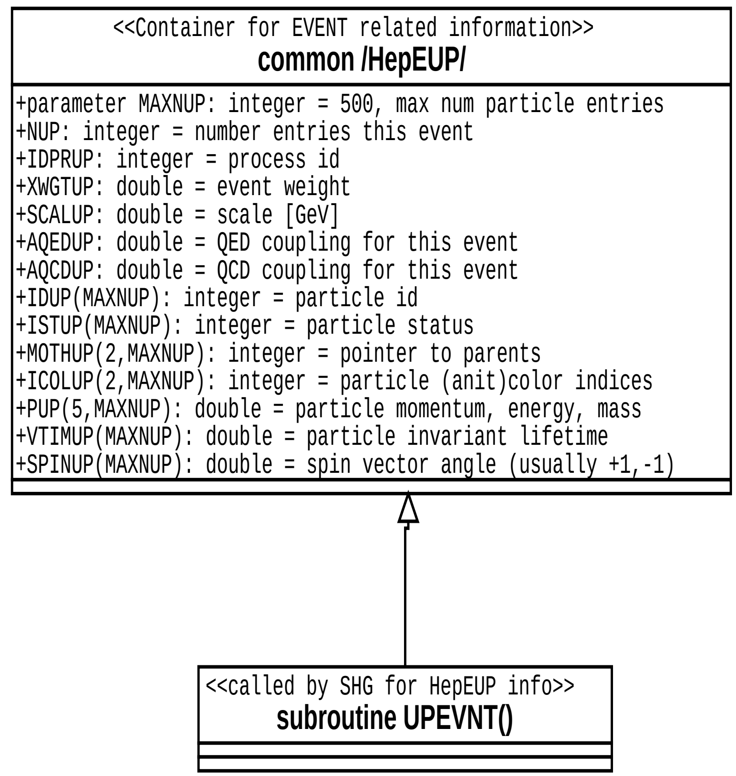
- Each stage (run and event associated with own subroutine)
- Run subroutine
  - IDWTUP: master switch indicating how the event weights (XWGTUP) are interpreted (some examples below)
    - +1: events are weighted on input and SHG is asked to produce events with weight +1 on output
    - -1: same as above but event weights may be either positive or negative; SHG will produce events with weights +1 or -1 on output
    - +3: events are unweighted on input so SHG only asks for next event
    - -3: same as above but event weights may be either +1 or -1



(Specialized for

Event related information

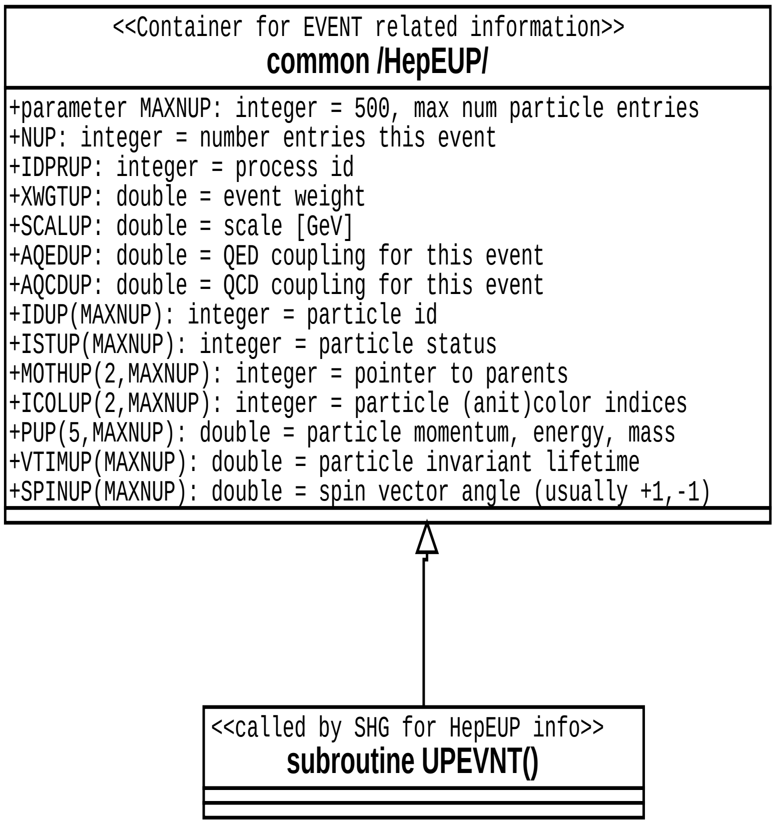
- NUP: number of particle entries for this event
- IDPRUP: ID of the process for this event
- XWGTUP: event weight
- IDUP: particle ID (non-physical particles assigned IDUP=0)
- ISTUP: status code
  - -1: incoming particle
  - +1: outgoing particle
  - -2: intermediate space-like propagator defining an x and Q^2 which should be preserved (DIS-specific)
  - +2: intermediate resonance, mass should be preserved
    - recoil from parton shower needs to be absorbed by particles in the event
  - +3: intermediate resonance, for documentation only
  - -9: incoming beam particles



each matrix element)

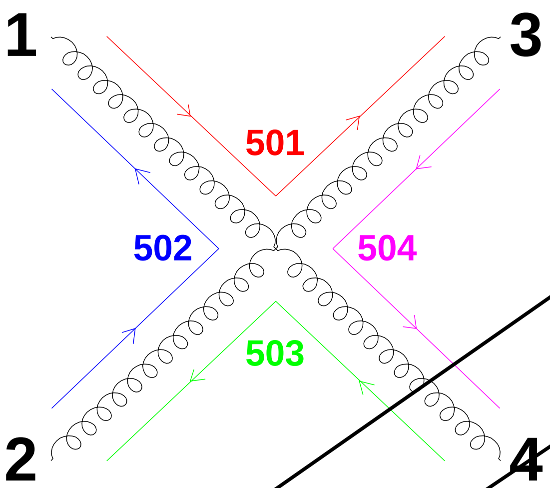
Event info

- MOTHUP(2,I): index of first and last mother
  - For decays, daughter particles will only have 1 mother
  - For 2->n, daughter particles will have 2 mothers
- Color flow: specific choice of color flow for a particular event is often unphysical, due to interference effects, but SHGs require specific color state from which to begin shower
  - ICOLUP(1,I): integer tag for color flow line passing through color of the particle
  - Integer tag fro color flow line passing through anti-color of tag



each matrix element)

Example (gg->gg)



initial/final state

particle code

mother/daughter relationships

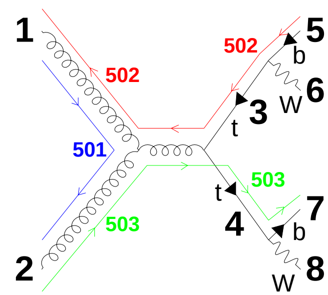
color flow

I	ISTUP(I)	IDUP(I)	MOTHUP(1,I)	MOTHUP(2,I)	ICOLUP(1,I)	ICOLUP(2,I)
1	−1	21 (g)			501	502
2	−1	21 (g)			502	503
3	+1	21 (g)	1	2	501	504
4	+1	21 (g)	1	2	504	503

Consider ttbar production

- t and tbar given ISTUP=+2, which informs SHG to preserve their invariant masses when showering and hadronizing the event
- Intermediate s-channel gluon has been drawn, but no entry because cannot be distinguished from t-channel
- Definition of color or anti-color line depends on orientation of graph
  - define color and anti-color according to physical time order
  - quark will always have color tag ICOLUP(1,I) filled, but never its anti-color tag ICOLUP(2,I); reverse for anti-quark; gluon has info in both tags

Example: hadronic $t\bar{t}$ production



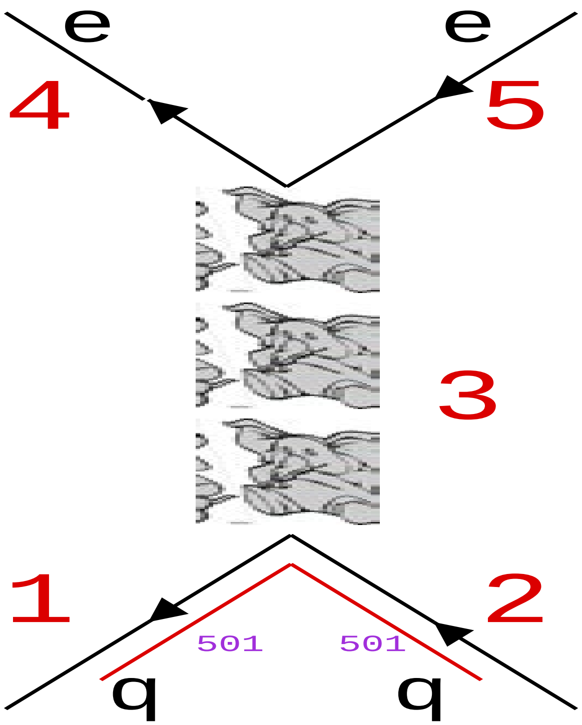
I	ISTUP(I)	IDUP(I)	MOTHUP(1,I)	MOTHUP(2,I)	ICOLUP(1,I)	ICOLUP(2,I)
1	−1	21 (g)	0	0	501	502
2	−1	21 (g)	0	0	503	501
3	+2	−6 ($\bar{t}$)	1	2	0	502
4	+2	6 (t)	1	2	503	0
5	+1	−5 ($\bar{b}$)	3	3	0	502
6	+1	−24 (W^-)	3	3	0	0
7	+1	5 (b)	4	4	503	0
8	+1	24 (W^+)	4	4	0	0

The t and $\bar{t}$ are given ISTUP=+2, which informs the SHG to preserve their invariant masses when showering and hadronizing the event. An intermediate s-channel gluon has been drawn in the diagram, but since this graph cannot be usefully distinguished from the one with a t-channel top exchange, an entry has not been included for it in the event record.

The definition of a line as 'color' or 'anti-color' depends on the orientation of the graph. This ambiguity is resolved by defining color and anti-color according to the physical time order. A quark will always have its color tag ICOLUP(1,I) filled, but never its anti-color tag ICOLUP(2,I). The reverse is true for an anti-quark, and a gluon will always have information in both ICOLUP(1,I) and ICOLUP(2,I) tags.

Note the difference in the treatment by the parton shower of the above example, and an identical final state, where the intermediate particles are not specified:

Another example:
little pink elephant exchange



I	ISTUP(I)	IDUP(I)	MOTHUP(1,I)	MOTHUP(2,I)	ICOLUP(1,I)	ICOLU
1	−1	−2 ($\bar{u}$)	0	0	0	5(
2	−1	2 (u)	0	0	501	(
3	+2	0 (pink elephant)	1	2	0	(
4	+1	11 (e^-)	3	3	0	(
5	+1	−11 (e^+)	3	3	0	(

Effective use of pdf uncertainties

- PDF uncertainties are important both for precision measurements (W/Z cross sections) as well as for studies of potential new physics (a la jet cross sections at high E_T)
- Most Monte Carlo/matrix element programs have "central" pdf's built in, or can easily interface to PDFLIB
- Determining the pdf uncertainty for a particular cross section/distribution might require the use of many pdf's
  - CTEQ Hessian pdf errors require using 33 pdf's
  - GKK on the order of 100
- Too clumsy to attempt to includes grids for calculation of all of these pdf's with the MC programs
- **→Les Houches accord #2**
  - each pdf can be specified by a few lines of information, if MC programs can perform the evolution
  - fast evolution routine will be included in new releases to construct grids for each pdf
- NB: pdf uncertainties make most sense in the context of NLO calculations; current MC programs are basically leading order and LO pdfs should be used when available
  - NNB: CTEQ6L is a leading order fit to the data but using the 2-loop α_s, since some higher order corrections are in MC programs like Pythia, Herwig, etc

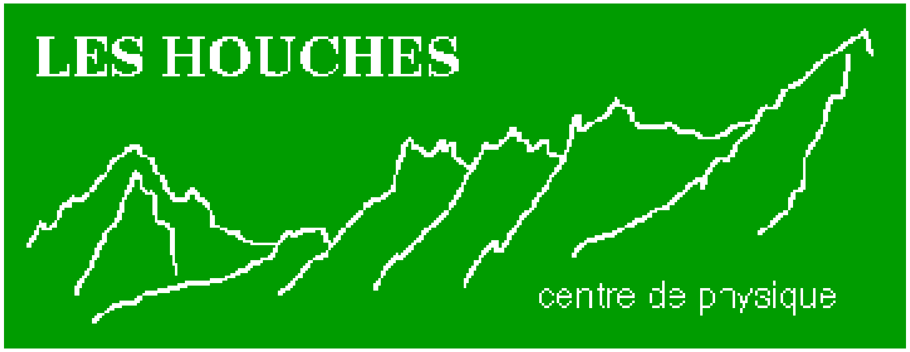
Les Houches accord #2

- Using the interface is as easy as using PDFLIB (and much easier to update)
- First version has CTEQ6M, CTEQ6L, all of CTEQ6 error pdfs and MRST2001 pdfs
- See pdf.fnal.gov (and talk by Walter Giele at this conference)

- call InitiPDFset(*name*)
  - called once at the beginning of the code; *name* is the file name of external PDF file that defines PDF set
- call InitPDF(*mem*)
  - *mem* specifies individual member of pdf set
- call evolvePDF(*x*,*Q*,*f*)
  - returns pdf momentum densities for flavor *f* at momentum fraction *x* and scale *Q*

The Big Idea

- Reminder: the big idea:
  - The Les Houches accords will be implemented in all ME/MC programs that experimentalists/theorists use
  - They will make it easy to generate the multi-parton final states crucial to much of the Run 2/HERA/LHC physics program and to compare the results from different programs
  - experimentalists/theorists can all share common MC *data sets*
  - They will make it possible to generate the pdf uncertainties for any cross sections

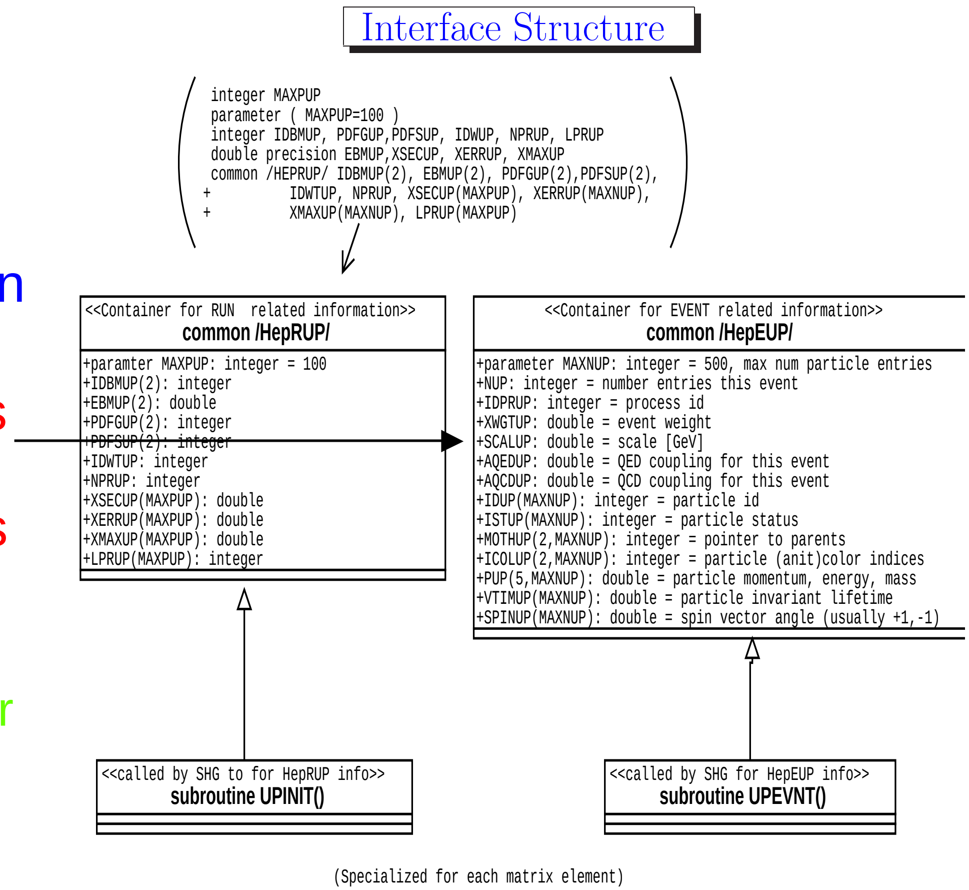


Les Houches accords

- Les Houches accord #1 (ME->MC)
  - accord implemented in Pythia 6.2
  - accord implemented in Herwig 6.5
  - accord implemented in CompHEP
    - CDF top dilepton group has been generating ttbar events with CompHEP/Madgraph + Pythia
  - accord implemented in ALPGEN
    - hep-ph/0206293
  - accord implemented in Madgraph
    - MADCUP:http://pheno.physics.wisc.edu/Software/MadCUP/}.
    - MADGRAPH 2: within a few weeks
  - Implemented in Grace
  - in AcerMC:hep-ph/0201302
- Les Houches accord #2 (pdfs in ME/MC)
  - version of pdf interface has been developed
    - available at http://pdf.fnal.gov
  - commitment for being implemented in MCFM
  - commitment for being implemented in *your name here*

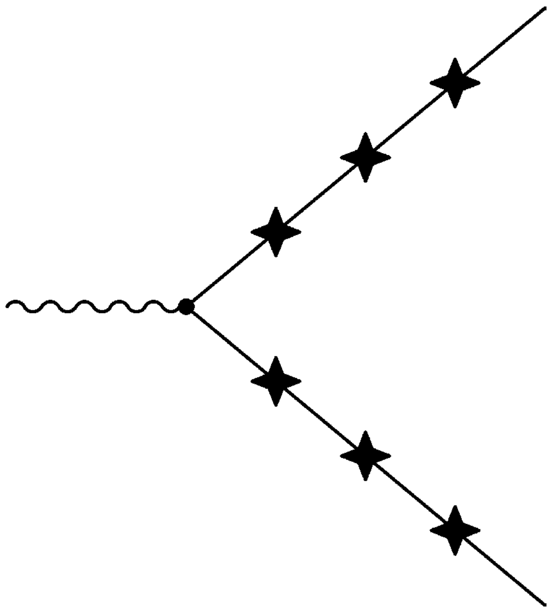
What Les Houches doesn't do

- Specify the exact form of output format
  - nominally details are supposed to be invisible to *casual* user
- Specify *correct* Q scale for parton showering
  - imagine that W + 5 jet events probe more deeply into shower than W + 1 jet events
    - would need lower scale
    - has to be provided by hand for Pythia; Herwig uses color flow
- Provide a seamless flow from matrix element to parton shower
  - correct Sudakov form factor



The proper way

- The proper way of taking care of this would be to generate parton showers starting at full hard scale but vetoing those emissions that populate same phase space as exact ME
  - see for example, Krauss, Catani, Kuhn and Webber, hep-ph/00109231
- Frank is working on an implementation of this procedure in a hadron-hadron Monte Carlo
- For the moment (Winter conferences), need to look for the *next best solution*
  - this workshop can help to provide a better understanding of the issues



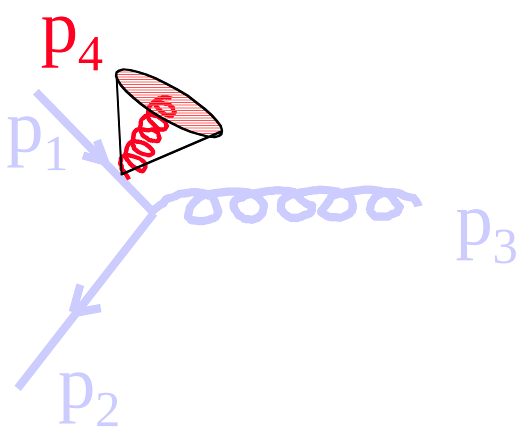
Michelangelo's prescription

- CDF (and D0) are preparing Monte Carlo samples of W/Z + jets for the Winter conferences
  - necessary to understand W/Z + jets(/heavy flavors) as backgrounds to top/Higgs/new physics
    - a good understanding of QCD production mechanisms is even more important now since b-tagging tools are not fully developed
  - using ALPGEN/MADGRAPH/CompHEP/ GR@PPA for matrix element generation and Herwig and Pythia for parton showering and hadronization
    - one of the first steps is to see if all ME programs give the same result with the same input parameters/cuts (see talk by Gervasio)
  - what cuts/parameters should be used for the matrix element generation?
    - how can double-counting/under-counting be avoided?

What systematics (for ME/MC generation)??

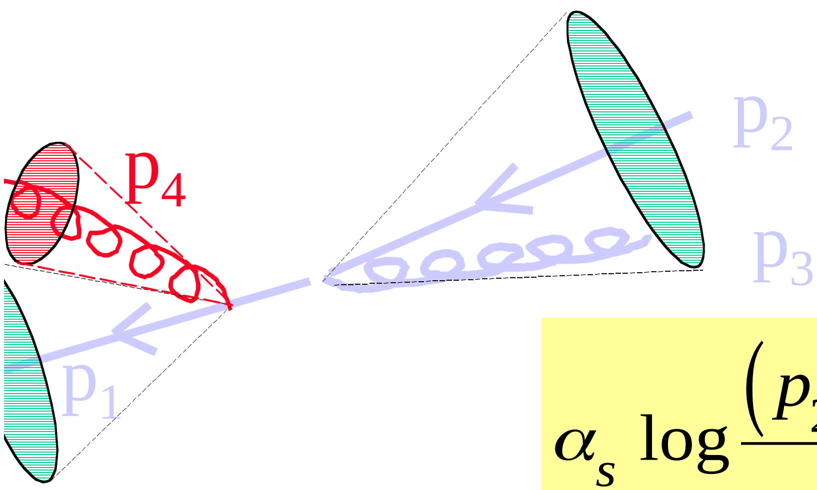
- In the generation of parton level samples to be processed through a shower evolution, we need to keep the *parton-level cuts not harder than the jet cuts*, else we lose the advantage of the correct description of hard, large-angle emission by the ME calculation
- So a reasonable starting point is to set
  - $p_{T\,parton} > p_{T\,min} = E_{T\,jet}^{min}$ *and* $\Delta R(\text{parton-parton}) > \Delta R_{cut} = \Delta R_{jet}$
- However these thresholds may not be sufficient to guarantee full generation efficiency. Parton configurations not passing these cuts might still give rise to hadronic final states passing the final jet cuts. For example, a jet below threshold might be pushed above thanks to some extra underlying event energy. As a result, one should start from softer parton-level cuts,
  - $p_{T\,min} < E_{T\,jet}^{min}$ *and* $\Delta R_{cut} < \Delta R_{jet}$
- A good, stable, *parton → shower merging algorithm* would give jet X-sections which, aside form the "efficiency effects" mentioned earlier, should be independent of the parton-level generation cuts, and in particular should converge to a finite answer for *and* $p_{T\,min} \rightarrow 0$ $\Delta R_{cut} \rightarrow 0$. The X-section should only depend on the jet-level cuts (ΔR_{jet} *and* E_{Tjet})

Leading vs subleading double counting
Example: corrections to 3-parton final states



is of $O(\alpha_s)$ relative to the LO process

unless:



which gives a contribution to $\sigma_{3\text{-jet}}$ of order

$$\alpha_s \log \frac{(p_2 + p_3)^2}{E^2_{T\ jet}} \approx \alpha_s \left(\log \frac{p_T^{\max}}{p_T^{\min}} + \log \frac{1}{\Delta R} \right)$$

Double counting is sub-leading provided ΔR and are not too large

$p_T^{\max} / p_T^{\min}$

Progress towards solutions (II) vetoed showers

(Catani, Krauss, Kuhn, Webber)->see also Steve's talk at last tuning worksho

- Generate samples of different jet multiplicities according to exact tree-level ME's, with N_{jet} defined using a k_{perp} algorithm

$$y_{ij} = \frac{2\min\left\{E_i^2, E_j^2\right\}\left(1-\cos\theta_{ij}\right)}{s} \geq y_{cut} = \frac{Q_{cut}^2}{s}$$

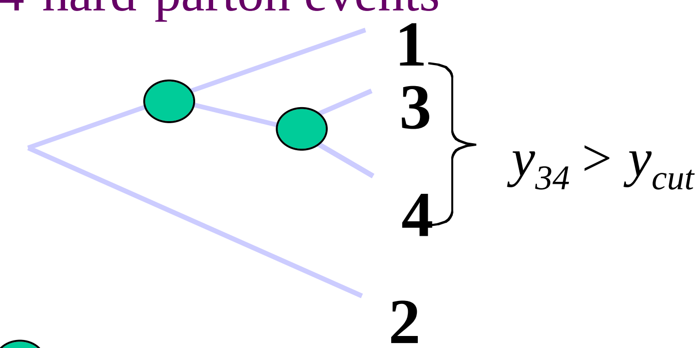
Reweight the matrix elements by vertex Sudakov form factors, assuming jet clustering sequence lefines the colour flow->

Remove double counting by vetoing shower histories (i.e. y_{ij} sequences already generated by the natrix elements)

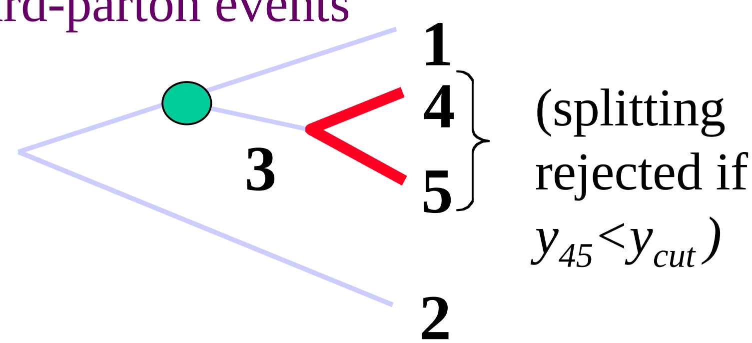
This procedure takes into account the full ME for jet production at tree level plus all HO eading and next-to-leading logarithmic contributions.

Fully successfull for e^+e^- collisions, being extended to hadronic collisions (see note later)

From the sample of 4-hard-parton events

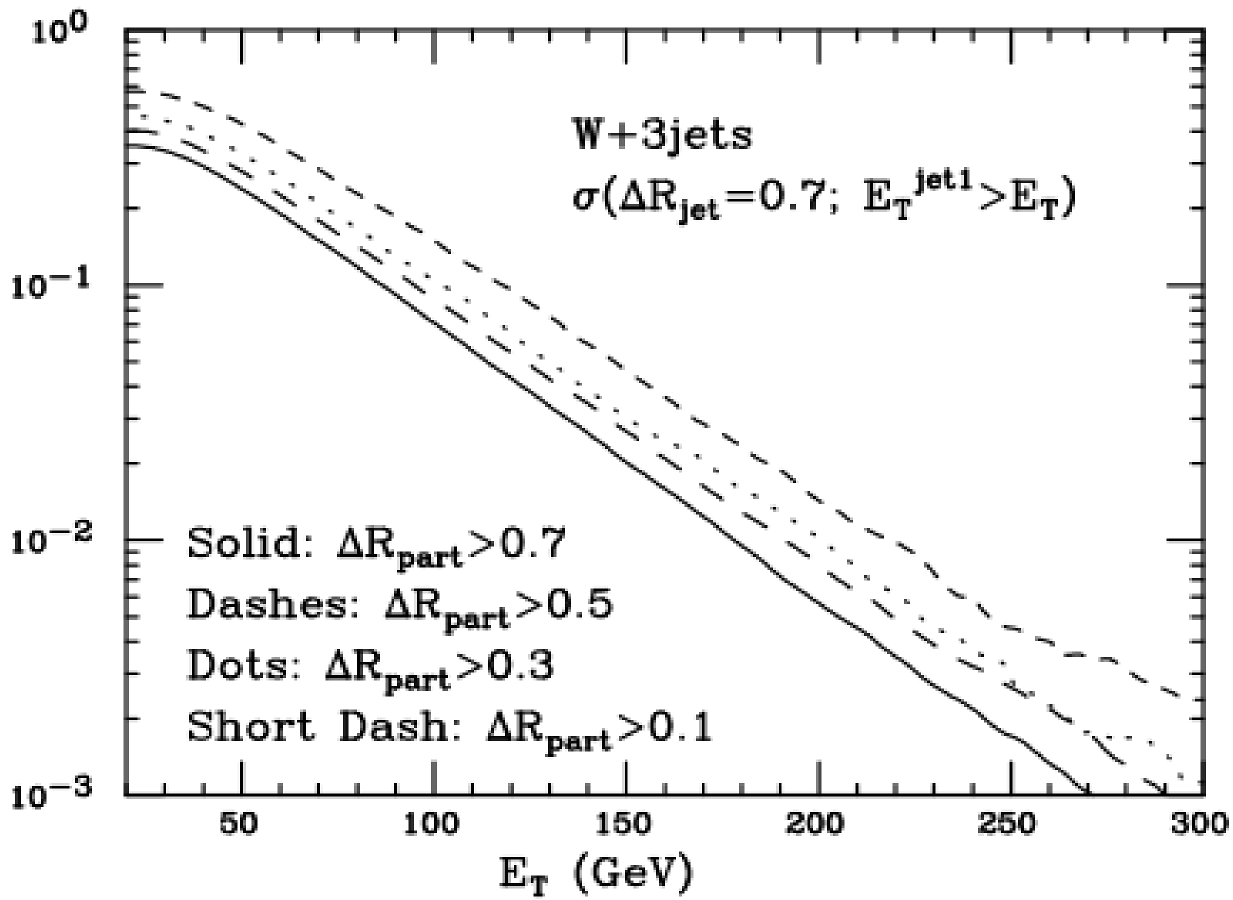


From the sample of 3-hard-parton events



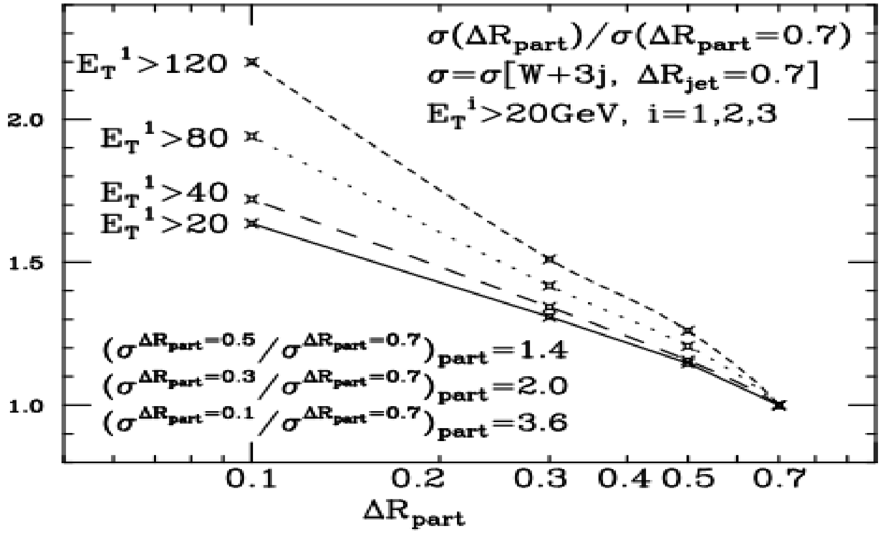
: Sudakov correction

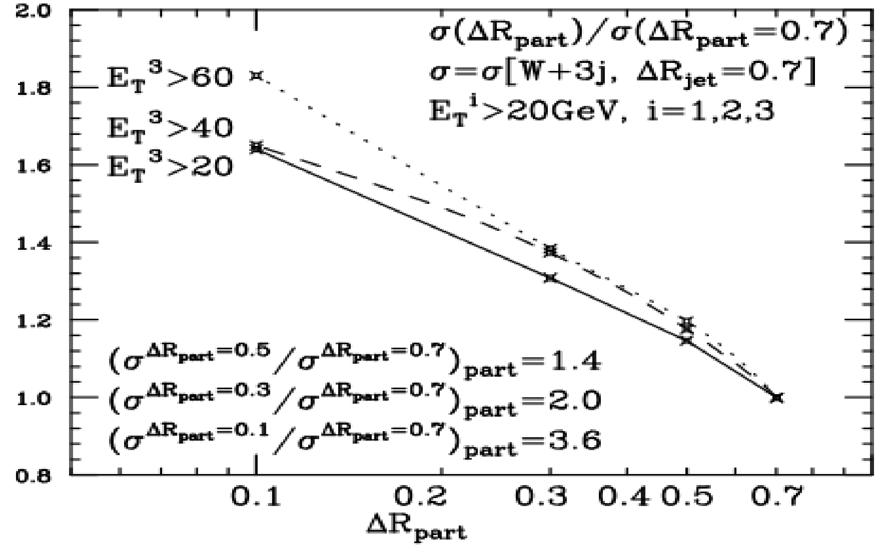
Spectrum of the leading-E_T jet (jet1)

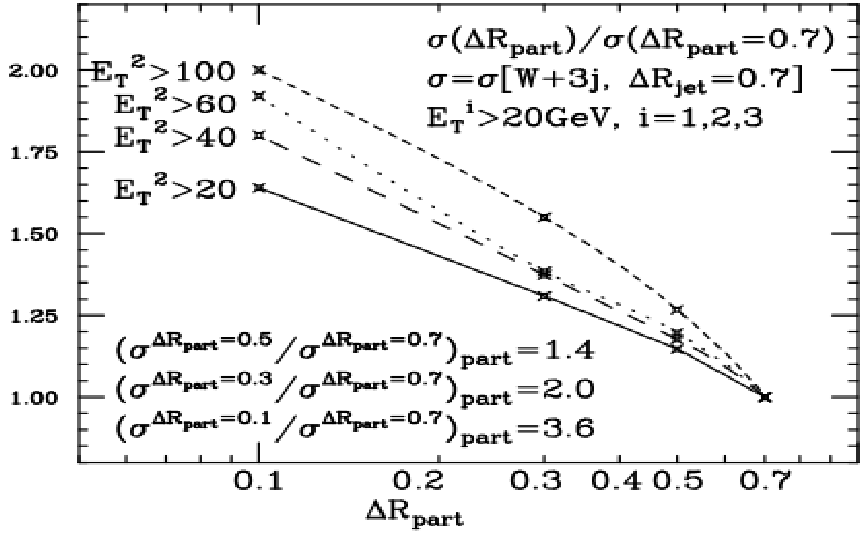


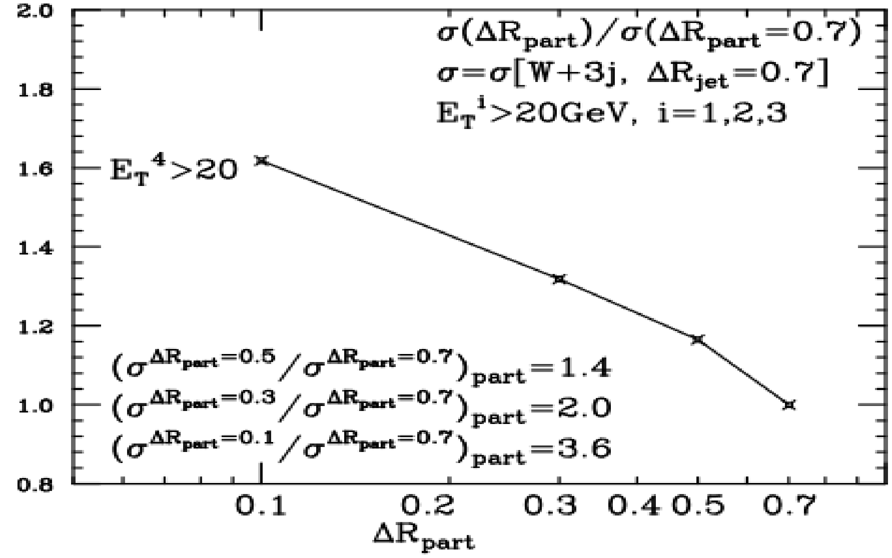
The ΔR_{part} dependence becomes more significant at high E_T, as expected because of larger logs

E_T spectrum dependence on ***ΔR_{part}*, for the 4 most energetic jets**









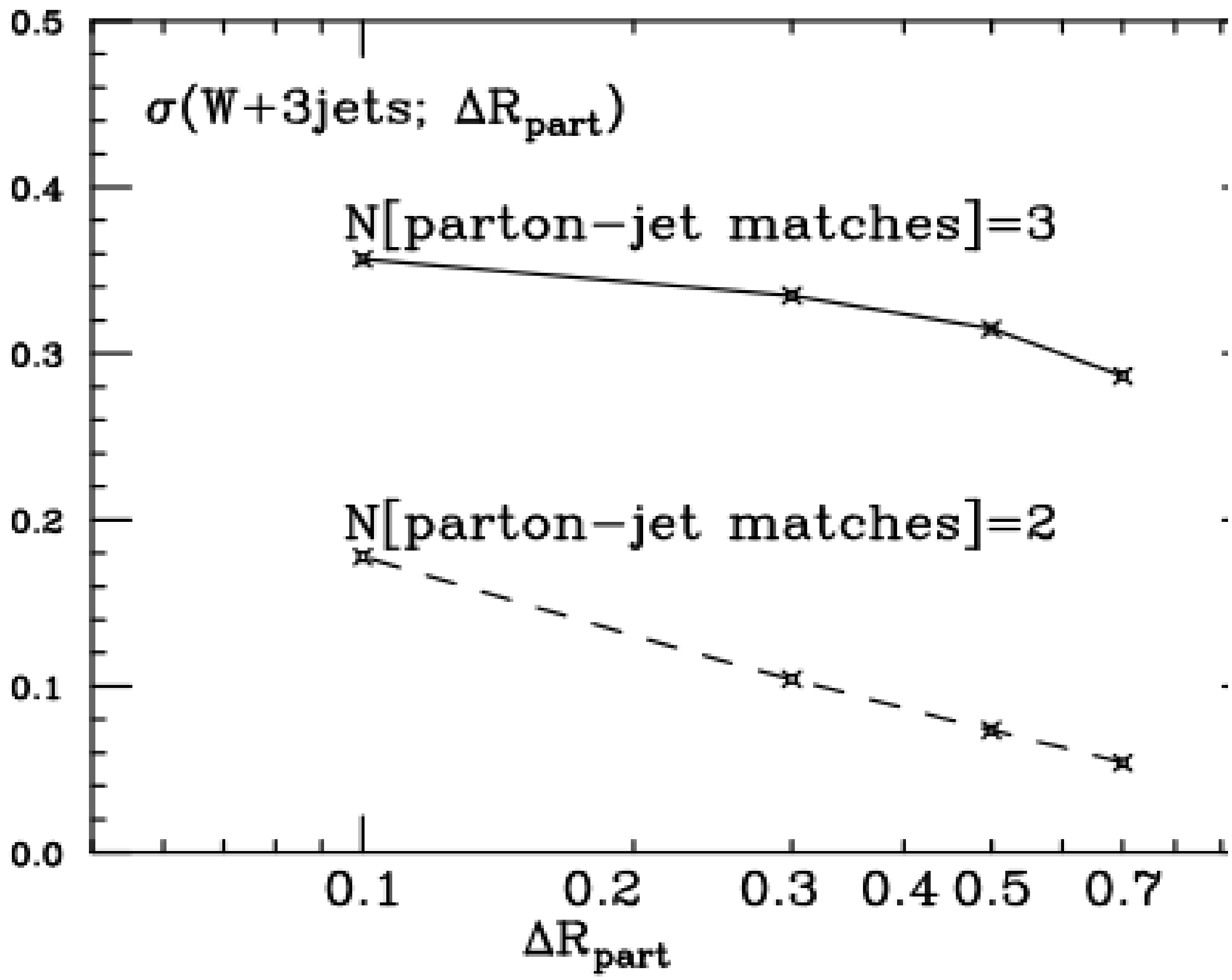
Much larger sensitivity than in the case of *$p_{T\,min}$*

Matching partons and jets

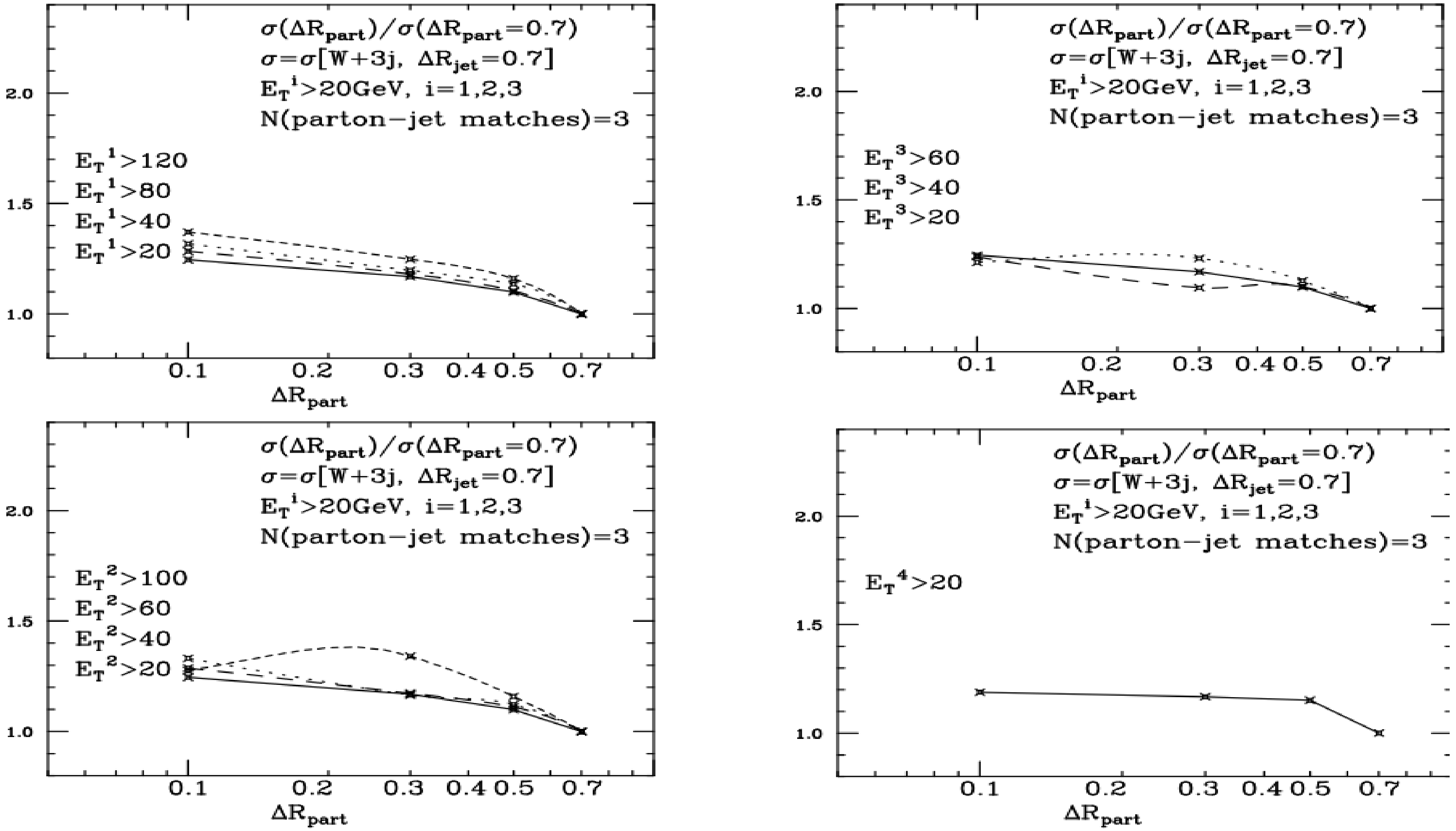
latching criterion:

=#(parton-jet) pairs
ith **ΔR<** 0.7
only 1 jet can be
ssigned to a given
arton)

***he rate for events
ith all partons
atched by a jet is
ather flat, and
aturates as***
R_{part}->0



Cone dependence of Et distributions for events with N_{match}=3



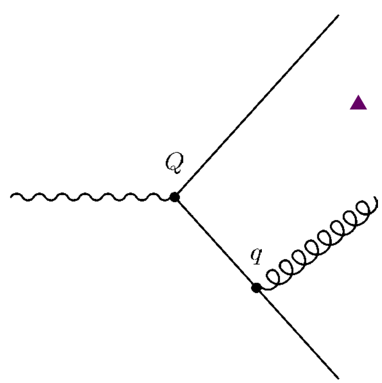
Michelangelo's marching orders

Take a W+n parton event. Let us shower it, and apply "some" jet clustering algorithm (ideally directly on the shower-MC output, not on the output of the detector simulation). This will lead to N jets. If N<n, we throw away the event. If N>=n, we proceed to "matching":

```
nmatch=0
do iparton=1,n
     do ijet=1,N
          if (iparton.MATCHES.ijet .
              and. ijet-has-a-match-
              already.eq.false ) then
                nmatch=nmatch+1
                ijet-has-a-match-
              already=true
          endif
     enddo
enddo
```

For younger viewers,
this is Fortran

- **This is a prescription, not a theorem**
- What may be wrong with this prescription?
  - radiation dip at $Q=Q_{jet}$?
    - **homework assignment: check to see how large of an effect this really is**
    - rate for 2 jets at scale Q_{jet} is:
      $R_2^{NLL}(Q_{jet})=[\Delta_q^{NLL}(E_{cm},Q_{jet})]^2$
    - correct rate for 2 jets at lower scale Qo is:
      $R_2^{NLL}(Q_{jet})=[\Delta_q^{NLL}(E_{cm},Q_o)]2$
    - if parton shower is started at scale Q_{jet}, then rate becomes:
      $R_2^{NLL}(Q_{jet})=[\Delta_q^{NLL}(E_{cm},Q_{jet})\Delta_q^{NLL}(Q_{jet},Q_o)]^2$
    - can recover the correct rate if start parton shower at E_{cm}, but veto emissions with $q>Q_{jet}$



Scales

- Can we jigger scale to account for this? Do we need to?
  - scale is provided by Les Houches accord to Pythia; easier to change
  - Herwig knows the scale from color flow; harder to change
  - if use a larger scale, do we now double-count without a parton shower veto
    - **homework assignment: check if parton shower produces jets at same or greater y scale (k_T algorithm) as ME does; this would be double-counting**

What are we doing now?

- Generate W + n jet events with $\Delta R = 0.2$, p_{T}_min = 8 GeV/c (for jet size R=0.4, p_{T_min}=15 GeV/c)
- Feed through Herwig/Pythia showering/hadronization
- Apply Michelangelo's matching scheme to jets in event (see talks by Gervasio,Andrea for more details)
  - **homework assigment: check Michelangelo's results for ΔR, p_{T}_min sensitivity at detector level**
  - use clustering machinery developed by Matthias Tonnesmann

    does it matter whether at parton level/hadron level?

    homework assignment: Check.
  - classify event as < n jet, n jet, > n jet
    - throw away < n jet events
    - throw n jet events into exclusive sample pile
    - throw n jet, > n jet events in to inclusive sample pile
  - exclusive sample is obtained by Σ-ing over the n exclusive piles
  - inclusive samples are good for >= n jets in final state
- How well does the parton showering produce > n jets?
  - we have some experience with W + 1 jet ->W + 2,3,4 jets in Run 1 (increasingly bad)
  - conjecture: problems are less severe if starting n is large (3,4); simplicity then in generating large n final states
    - **homework assignment: check**

Short/medium term

- Use NLO calculations to check prescription
  - **homework: compare ME W + 2 jets (+ parton showers) to NLO W + 2 jets from MCFM**
    - what comparisons would be useful?
    - can we learn anything about best scales, ΔR, p_{Tmin} cuts to apply to generation of ME->MC events?
- **Homework: Compare to existing data**

Longer term

- Frank Krauss and students are working on a new parton showering MC program which incorporates full SM matrix elements a la Madgraph (MSSM in preparation), with parton shower vetoing as discussed earlier; beta release in spring
- I suggested to him that CDF/D0 people may be reluctant to use a completely new program until fully vetted but that his technique could be well applied to Herwig and/or Pythia
  - his response was that his going back to using Fortran would be like CDF returning to using a bubble chamber
  - maybe Steve will implement processes in Fortran Pythia
  - in longer term, I think there is interest in doing this type of tree level ME correction in Herwig (in C++ release)
  - MC@NLO has this correction implicitly for 3 parton final states

Conclusions

- It would be nice not only to collect talks on a website but also to summarize our discussion/conclusion in a note
- Lastly, remember, this is leading order physics